A SOCIO-ECONOMIC TAXONOMY OF ARAB COUNTRIES

By

Imed Limam

January 1998

ABSTRACT

Many classifications of Arab countries are based on arbitrary criteria and on a subjective delimitation of groups that rely on commonly used stereotypes. These classifications have also neglected several important socio-economic characteristics of the Arab countries.

The paper attempts to adopt an overall as well as sector-specific classifications of Arab countries based on a wide range of socio-economic variables, over three time periods, using different multivariate statistical methods for classification. The results show that "oil-versus-non-oil" represents the most meaningful criterion to adopt overall as well as sector-specific classifications of the Arab countries. Income level and the degree of diversification of the economies are found to be of some relevance in certain sector-specific classifications. The results also reveal some important taxonomies of Arab economies. The most obvious is that the six countries of the Gulf Cooperation Council (GCC) represent the most homogeneous group of Arab countries with respect to the main socio-economic variables used in the paper.

1. INTRODUCTION

The literature abounds with classifications of countries for many fundamental purposes. The general objective of classification is to group countries into sub-groups that differ in a meaningful way with respect to certain criteria.

Classification analysis can be used in a wide variety of disciplines. For instance, in Psychology it can help classify persons according to personality types, in Urban Planning it can classify cities according to demographic criteria, and in marketing, it can be used to classify customers according to taste.

One of the widely used applications in Economics, is to classify countries according to their level of development. The World Bank (WB) and the United Nations (UN) are among the international institutions interested in this type of classifications. In WB (1996), countries are classified into four income groups (GNP per capita) and five regions.[1] The income groups are: low-income ($ 765 or less), lower-middle income ($ 766- $ 3,035), upper-middle income ($ 3,036- $ 9,385), and high-income ($ 9,385 or more). The five regions being: Sub-Saharan Africa, Asia, Europe and Central Asia, Middle-East and North Africa, and the Americas. In the same document, the WB also adopts a classification of countries by export category and indebtedness.

The United Nations Development Program (UNDP) classifies countries according to an index composed of three indicators namely, life expectancy, education, and income.[2] The index is dubbed the Human Development Index (HDI). The classification of countries according to this index was published in the Human Development Report of UNDP in 1991. This report distinguishes between three groups of countries according to their HDI scores.[3] Countries with an index equal or above 0.8 are classified as "High Human Development" countries. If HDI is between 0.5 and 0.799, the countries are classified as "Medium-Human Development" countries. The group of "Low Human Development" comprises countries with an HDI score that is less than 0.5.

[1] See for instance, World Bank (1996), pp. 576-77.

[2] See UNDP(1990, 1991), for a precise definition and analysis of the index.

[3] UNDP (1991), p. 197.

Many institutions and studies have tried to find a typology of Arab Countries according to different criteria.[4] The 1991 issue of UNDP's World Report has, for instance, classified Arab countries into three groups.[5] The first group incorporates oil-rich countries with fairly high human development level. The countries of the group are characterized by a lag between the level of their human development and their GNP rankings.

In contrast, the second group is characterized by a relative balance between their economic and human development. The group includes countries like Tunisia, Syria, Lebanon, and Jordan. These countries have been successful in achieving above average human development in the region and an even distribution of income.

The third group, incorporates countries with low levels of income such as Sudan and Yemen.

The 1997 issue of the Human Development Report based on 1994 estimates, shows that the classification of Arab countries according to HDI ranked Bahrain, U.A.E, Kuwait, Qatar, and Libya as countries with high human development; Lebanon, Saudi Arabia, Syria, Tunisia, Algeria, Jordan, Oman, Egypt, Morocco, and Iraq, as countries with medium human development; and the rest of Arab countries: Comoros, Yemen, Mauritania, Sudan, and Djibouti, as countries with low human development level.

The Unified Arab Economic Report, published by a group of four Arab institutions namely, the Arab Fund for Economic and Social Development (AFESD), the Arab Monetary Fund (AMF), the Organization of Arab Petroleum Exporting Countries (OAPEC), and the Arab League General Secretariat, used to classify Arab countries into oil-producing and non-oil producing. This classification has been avoided inasmuch as possible in recent reports following the discontent of many quarters with this criterion.

In a study conducted by OAPEC and the Italian Enter Nazionale Idrocarburi (ENI), to analyze interdependence among Arab and OECD economies, the Arab countries in the study were classified into three groups.[6] The general criteria used for classification were oil abundance, labor market characteristics, and the external sector.

[4] See Al-Kawaz (1995) for a brief review of some of the widely used classifications.

[5] UNDP (1991), p. 35.

[6] OAPEC and ENI (1985).

The first group consists of countries with large populations and substantial deficits in their balance of payments. These were: Algeria, Egypt, Iraq, and Syria. The second group consists of oil-producing countries with labor supply shortages namely, Bahrain, Kuwait, Libya, Qatar, Saudi Arabia, and the United Arab Emirates (UAE). The third group consists of oil-importing countries with large balance of trade deficits. These are: Jordan, Morocco, Tunisia, Yemen, and Oman.

In a project undertaken by the Centre for Arab Unity Studies, between 1981 and 1987, to analyze and compare possible future development scenarios for the Arab countries under study, a four-group classification was adopted.[7] The First group consisted of the six countries of the Gulf Cooperation Council (GCC), Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and UAE plus the two parts of Yemen at that time. The countries in this group have oil as the main natural resource, do not have diversified economies, have a marginal agricultural sector, small populations, and predominantly unskilled labor force that tends to be concentrated in the service sector.

The second group, dubbed Al-Mashriq, comprises Syria, Lebanon, Jordan, and Iraq. The countries of this group have in common, high agricultural and industrial capacities, as well as a multi-ethnic social fabric.

The third group is made of four Maghreb countries: Algeria, Mauritania, Morocco, and Tunisia. These countries have diversified economies, large bases of natural and human resources, and are financially constrained and debt-ridden.

The final group labeled the Nile group, consists of Egypt, Libya, and Sudan. Somalia and Djibouti can also be linked to this last group. These countries have fertile arable lands, a great agricultural potential, and weak industrial and infrastructure bases.

In an individual study, Amrouch and Wadie (1987) have classified Arab countries based on 58 indicators related to the economic structure and level of development, demography and population, education, and labor market characteristics.

Using multivariate statistical analysis, the authors have identified the most important factors accounting for most of the total variance in the variables under study. Then, they used the score of each country with respect to the latter factors to classify them using hierarchical clustering techniques.

[7] See for instance, Saadeddine E. et al. (1989), and Hassib, K. et al. (1988).

Based on their results, the authors have classified the twenty one Arab countries under study into six groups. Group 1 comprises countries with high reliance on oil as a source of income, similar population densities, and high school enrollment ratios. These countries are: Bahrain, Qatar, and UAE.

Group 2, contains only Djibouti characterized by its small population size, high reliance on the service sector, high urbanization rate, and low mortality rate.

Six countries, are included in group 3. These are Oman, Saudi Arabia, Libya, Jordan, Iraq, and Kuwait. The countries of the group have a fairly close level of income and population growth rates, rely on extracting industries, and have high levels of education and life expectancy.

Group 4 is made of Algeria, Egypt, Morocco, Syria, and Tunisia. They share in common the same economic structure and several demographic and educational characteristics. Lebanon makes up group 5 with its high reliance on the service sector and high level of income, life expectancy, enrollment ratios, and urbanization.

Finally, group 6 is made of the poorer Arab countries characterized by a heavy reliance on the agricultural sector and low scores in human development indicators. These countries are the two parts of Yemen, Somalia, Sudan, and Mauritania.

Most of the known classifications of Arab countries are based on arbitrary criteria and on a subjective delimitation of groups that relies heavily on general factors such as income, natural or human resource endowments, and geographical contiguity. Many other socio-economic characteristics of Arab countries have been neglected. Moreover, the purpose of these classifications, and especially those emanating from international institutions, has been very often the ranking of the countries according to their development level.

Classification of countries according to their degree of similarities with respect to various characteristics, is also extremely important for international and regional donor and development institutions monitoring the socio-economic situations in their member countries. Such classification facilitates the analysis by making cross-country comparisons, extrapolations, and generalizations possible.

In this paper, I will attempt to adopt an over all as well as sector - specific classifications of Arab countries on the basis of a wide range of socio-economic characteristics using multivariate statistical methods. These classifications should not only provide objective typologies of Arab countries but also test for the validity of known stereotypes used previously in the classification of these countries. Finally, the outcomes of the paper should provide additional criteria for the classification of Arab countries for different purposes of analysis.

Section 2 rationalizes and underlines the methods used in the paper. Section 3 provides the catalogue of the indicators used in the analysis as well as data sources. Section 4 analyzes the results of the study. Finally, section 5 concludes.

2. METHODOLOGY

Multivariate Statistical Analysis has been widely used in classifying countries according to various criteria, notably their level of development. Several classifying techniques can be used.[8]

Discriminant Analysis classifies individuals or objects into mutually exclusive groups on the basis of a set of variables and an a priori available classification. Each subject is assigned a score which represents a weighed average of the subject's values on the different variables. The scores are then be used to predict the probabilities of any subject to be in any of the groups under consideration. [9]

Cluster Analysis, unlike Discriminant Analysis, operates a classification of (n) subjects described by a set of (p) variables in the absence of an a priori classification. The classification of subjects is done on the basis of a given "Distance Measure" or "Optimality Criterion".[10]

Factor Analysis is yet, another technique that can usefully divide different subjects into a number of classes.[11] It is generally used to analyze common factors underlying the data. Factor Analyses can be Exploratory intended to identify unknown common factors characterizing the data, or Confirmatory intended to confirm the influence of a priori known common factors.

However, Factor Analysis can also be used, in the case of exploratory Factor Analysis, to underline the quantitative differences among subjects based on their factor scores on the identified factors.

In this paper, I will use Cluster Analysis and Exploratory Factor Analysis to classify Arab countries based on a set of socio-economic variables. In a first stage, Cluster and Factor Analysis are used to operate an overall classification of Arab countries. Then, Cluster Analysis is used to establish a sector-specific classification of these countries. In operating the above classifications, a sensitivity analysis is

---

[8] See for instance, Dillon and Goldstein (1984), for a good overview of the available classification techniques.

[9] See Dellaportas (1983) for a study using Discriminant Analysis in classifying countries according to their level of development.

[10] A good reference in this regard is Everitt (1993). For recent research work using Cluster Analysis in classifying counties according to development level see for instance, Vogel (1993).

[11] See Gorsuch (1983), as a comprehensive reference in Factor Analysis.

conducted in order to ensure robustness of conclusions over time and across classification methods.

2.1 Clustering Techniques

The objective of clustering techniques is the division of n subjects, in this case countries, into groups based on a given set of p characteristics. The division of subjects into clusters is performed using some similarity (dissimilarity) or distance measure. In the case of continuous variables, distance-type measures are used. Otherwise, matching-type measures are used.

Mainly, there are two types of clustering techniques: the Hierarchical Techniques and Partitioning Techniques. With Hierarchical Techniques, the classification of subjects into groups is irrevocable. Once a subject has been classified within one cluster, it cannot be relocated to another one. Hierarchical Techniques can themselves be classified into Agglomerative and Divisive.

Agglomerative Methods, start from the situation where each subject forms its own cluster. Then, the two closest subjects, based on a given definition of metric distance, are joined. In the following step, either a third subject joins the first group, or two other subjects form a separate group. In each new step the number of clusters is reduced until all subjects are grouped in only one cluster.

The Divisive Methods do exactly the opposite. In relative terms, Agglomerative Clustering Techniques are more widely used and built in most statistical computer packages.

In contrast with Hierarchical Techniques, Partitioning Techniques do not require that the allocation of subjects into clusters be irrevocable. Subjects are moved around from one cluster to another until an optimum is reached by a formal and predefined objective function.

Each group of Clustering Methods, whether Hierarchical or Partitioning, can use an array of techniques and use various dissimilarity or distance measures in order to classify subjects into groups.

Each technique and similarity measure has its merits and characteristics. No single technique or similarity measure could be claimed superior for all types of

cases.[12] Different clustering techniques and distance measures should be used in order to assess the stability of cluster solutions and robustness of the results.

In this paper, four clustering techniques are used. The first three, namely, Single Linkage, Complete Linkage, and Ward's Method, belong to the family of Hierarchical Clustering techniques. The fourth technique, K-Means, is part of the family of partitioning techniques. The choice of these methods has been based on their relative success in clustering subjects under different conditions and their availability on most statistical packages.

Given the nature of the data, the similarity measures that will be used are of the distance type. In order to test the sensitivity of the results to different distance measures, Euclidean and Pearsons distances will be used. First, I will go very briefly over the clustering techniques used. Then, I will give the formal definitions of the two distance measures used.

Single Linkage (Nearest-Neighbor Method)

In this method distance between two clusters is defined as that of the closest pair of subjects where the pair is formed by one subject from each cluster. This method is illustrated by figure 1 :

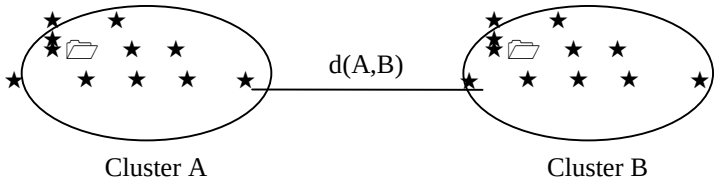

Figure 1 : Single Linkage Method

This method has been widely used in practice and was found to lead to more accurate description of the data than many other methods. However, Single Linkage Method has been found to cause chaining. Chaining occurs where subjects linked by a chain of intermediates are clustered together resulting in heterogeneous clusters. This method is not recommended when subjects are very similar.

[12] The relative merits of different clustering techniques and similarity measures are discussed in Everitt , and Dillon and Goldstein op. cit.

Complete Linkage (Furthest Neighbor Method)

The method defines distance in exactly the opposite way to the Single Linkage Methods. Distances between two clusters in the Complete Linkage Method is that between the most distant pair of subjects formed by one subject from each cluster. Figure 2 illustrates this method.

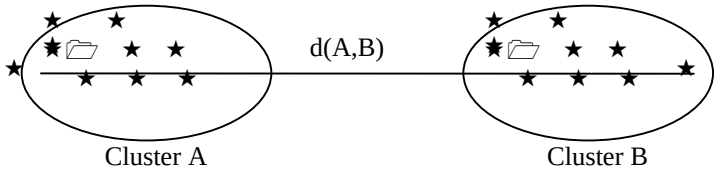

Figure 2 : Complete Linkage Method

Empirical investigations intended to give indicators as to the most useful clustering methods in practice, have found that the Complete Linkage Method performs more satisfactorily than many other methods. Complete Linkage shares a common quality with the Single Linkage Method namely, the invariance property. Invariance means that the clusters attained are invariant under monotonic transformation of the distance matrix. Furthermore, Complete Linkage Method is less sensitive to observational errors than several other methods and perform well in the case of unequal-sized clusters.

Ward's Method

In this method the distance between two clusters is given by the sum of squared deviations (SSD) between points and centroids. Let X_{ij} be the value of variable X for the i^{th} subject in the j^{th} cluster, k the total number of clusters and nj the number of subjects in the j^{th} cluster. SSD can be defined as follows :

$$SSD = \sum_{j=1}^{k}\left(\sum_{i=1}^{nj} X_{ij}^2 - \frac{1}{nj}\left(\sum_{i=1}^{nj} X_{ij}\right)^2\right)$$

Assignment of subjects into clusters is performed as follows: in the first stage, k groups of subjects are formed where each group comprises only one subject. In the next stage, the first group is formed by joining the two subjects whose group yields the smallest value of SSD for all possible clusters of size two. In the following stage, SSD is computed under two different cases. In the first case, each of the remaining subjects is added to the first group. In the second case, all possible size-two groups of unclustered subjects are formed. This process continues through successive stages. In each stage, the number of clusters is reduced. Ward's Method, has proved useful in the case of nearly equally-sized clusters.

K-Means

K-Means clustering method is a partitioning clustering technique based on optimizing a clustering criterion. This method consists of allocating n subjects into a predefined and fixed number of clusters, k. Subjects are moved around from one cluster to another until it is not possible to reduce the value of the objective function defined as follows :

$$E(n, k) = \sum_{i=1}^{n} D[i, l(i)]^2$$

where

$$D[i, l(i)] = \left(\left[\sum_{j=1}^{p} x_{ij} - \overline{x}_j(\ell)\right]^2\right)^{1/2}$$

is the Euclidean distance between subject i and the cluster mean for the variable j, $\overline{X}_j(\ell)$, of the cluster containing the i^{th} subject, l (i).

With respect to distance measures that can be used, many have been proposed in the literature.[13] Euclidean distance has been the most widely used measure. Let X_{ij} be the value of variable j for subject i, then the Euclidean distance between two subjects i and k is given by

$$d(i,k) = \sqrt{\sum_{j=1}^{P} (x_{ij} - x_{kj})^2}$$

Another widely used distance measure is the Pearson distance defined as follows :

$$d(i,k) = \sqrt{\sum_{j=1}^{P} (x_{ij} - x_{kj})^2 / v_j}$$

where v_j is the variance of variable j.

Many object the use of Pearson distance measure on the ground that it measures colinearity rather than similarity. Others have suggested its use because, it was able to pick up characteristics of the data not revealed by other distance measures.

Many distance measures are not scale invariant. In this case, it is customary to standardize the variable at hand. Standardization might lead to a loss of information and a dilution of differences among groups. However, it has the merit of preserving the relative distance among subjects.

2.2 Factor Analysis

Factor Analysis is a data reduction technique intended to summarize the interdependence among a given set of variables in a smaller number of factors that account for most of the common characteristics of the original variables.

[13] Note that the difference between distance measures: Euclidean and Pearson, is like the difference between inches and centimeters. These distance measures should not be confused with the methods of measuring the distance such as the Complete, Single, and Ward methods.

As indicated earlier, Factor Analysis is useful in underlying the common characteristics of the data as well as providing quantitative evidence about the difference between the subjects under study with respect to the identified factors.

In what follows, the main factor-analytic model that will be used in the paper is outlined.

The Model

Let $X = \left(X_1, X_{2,} \ldots X_p\right)$ be vector of p variables and $f = \left(f_1, f_2, \ldots, f_q\right)$ of q factors $q < p$. The basic factor model can be laid out as follows:

(1) $$X_i = \sum_{j=1}^{q} \lambda_{ij} \; f_j + e_i, \quad i = 1, \ldots, p$$

or more compactly as:

(2) $$X = \Lambda f + e$$

Where Λ is a p×q matrix of factor loadings and **e** is a p-dimensional vector of unobservable variables representing unique factors.

Under the assumption that the unique parts of the variables are uncorrelated with each other and with the common factors $\left(\text{cov}(e, f') = 0\right)$, the covariance of the vector of variables, X, can be written as follows:

(3) $$\sum = \Lambda \Phi \Lambda' + \Psi$$

where Φ i the covariance matrix of the common factors, and Ψ is the covariance matrix of the unique factors.

In order to facilitate the interpretation of the model, equation (3) can be written, without loss of generality, as follows:

(4) $$\text{var}(x_i) = \text{var}(c_i) + \text{var}(e_i) = h_i^2 + \psi_i \ , \ i = 1, \ldots, p$$

where h_i^2 is the common variance or communality of the variable X_i and Ψ_i is the variance of the unique factors. The communality of a variable is simply the portion of its total variance accounted for by the common factors.

It should be pointed out that if the common factors are uncorrelated and if the original variables are standardized (Sum of squares of the elements of the i^{th} row: contribution of all factors in the variance of variable i then the variance of each variable is given as follows:

(5) $$\text{var}(X_i) = \sum_{j}^{q} \lambda_{ij}^{\ 2} + \Psi_i \ , i = 1, \ldots, p$$

sum of squares of the elements of the j^{th} column: contribution of factor j in total variance of the entire set of variables.

where $\sum_{j=1}^{q} \lambda_{ij}^{\ 2}$ is the sum of squares of the ith row of the matrix Λ and represents the contribution of the different factors to the variance of X_i .

Symmetrically, the total contribution of factor f_j to the total variance of the whole set of variables, is nothing more than the eigenvalue of that same factor. It can be shown that this eigenvalue is the sum of squares of column j in Λ:

$$\underset{(pxq)}{\Lambda} = \begin{pmatrix} \lambda_{11} \cdots \lambda_{ij} \cdots \lambda_{1q} \\ \vdots \quad \vdots \quad \vdots \\ \lambda_{11} \cdots \lambda_{ij} \cdots \lambda_{1q} \\ \vdots \quad \vdots \quad \vdots \\ \lambda_{p1} \cdots \lambda_{pj} \cdots \lambda_{pq} \end{pmatrix}$$

If we define V_j as the contribution of factor f_j to the total variance of the entire set of variables, then it can be written as follows:

(6) $$V_j = \sum_{i=1}^{p} \lambda_{ij}^2 = \lambda_j' \lambda_j$$

We can also define the total contribution of all the factors to the total variance of all the variables as follows:

(7) $$V = \sum_{j=1}^{q} Vj = \sum_{i=1}^{p} \sum_{j=1}^{q} \lambda_{ij}^2$$

It follows that the relative contribution of factor j in total variance can be defined as follows:

(8) $$V_{cj} = \frac{V_j}{V}$$

It should be noted that the model presented so far can be rewritten in terms of correlation rather than covariance matrices. In fact, analyzing correlation is the basis of most factor-analytic models.

Let R be the correlation matrix of X, if the original variables are standardized then model (3) can be rewritten as:

(9) $$R = \Lambda \Phi \Lambda' + \Psi$$

$\Lambda \Phi \Lambda'$ is referred to as the common factor correlation matrix.

Factor analysis starts with the product moment correlation among variables and solve for the factor loadings , Λ, and factor correlation , Φ, given a set of appropriate restrictions. Factors are then extracted and factor scores computed.

As pointed out earlier, factor analysis amounts to looking at the amount of each variable's variance that is shared with other variables. Communalities are, therefore

central in factor analysis. There are several methods for computing communalities. Some of these methods use only few elements of the correlation matrix R and other methods use the entire correlation matrix.

Within the first type of methods, the communality of a variable i can be estimated for instance by the highest correlation of that variable with any of the remaining (p-1) variables, or by the average correlation of variable χ_i with the rest of the variables in the system. With regards to the second type of methods, the communality of a variable can be estimated by regressing the variable at hand with the rest of the variables in the system. The R^2 of this regression is then taken to be an estimate of the communality.

From equation (9), it is clear that the matrix that needs to be factored is not R but R^*, the reduced correlation matrix where the unit elements of the diagonal of R are replaced by the communalities of the respective variables. This follows from the need to estimate Ψ in order to solve for Λ and Φ.

In fact, one of the main factor solutions namely, the Principal Factor Method (PFM), uses the matrix R^* to extract factors as follows. PFM extracts factors such that "each factor accounts for the maximum possible amount of the variance in the set of the variables being factored".

This amounts to finding, first, the largest eigenvalue of matrix R^*. Let ∂_1 be that value and Y_1 the eigen vector associated with it, then it can be shown that:

(10) $\lambda_1 = \sqrt{\partial_1}\, Y_1$

The procedure continues by defining $R_1^* = R^* - \lambda_1\lambda_1'$ and finding λ_2 that maximizes $V_2 = \lambda_2'\lambda_2$ subject to the additional constraint that $R_1^* = R^* - \lambda_1\lambda_1' = \sum_{j=2}^{q} \lambda_j\lambda_j'$. [14] The procedure will not stop until all appropriate factors have been extracted. If we define R_m^* as the matrix resulting from subtracting the contribution of m factors, R_m^* can be written as follows:

[14] This implicity assumes that $\Phi = I$.

(11) $$R_m^* = R^* - \sum_{j=1}^{m} \lambda_j \lambda_j'$$

It follows that a good stopping rule is when R_m^* contains only small values. We can also define an equivalent stopping rule in terms of the eigenvalues of R^*. Since the sum of the eigenvalues of R^* is nothing more than the total communality V, then it makes sense to stop extracting factors when the sum of the eigenvalues is close to the value of the total communality.

Having extracted the q appropriate factors, one should be able to interpret the factor loading matrix in the sense of looking at which variables load into what factors. Then, a label is assigned to each factor. Such labels should depict the average meaning of the variables that load into each factor.

The common factors thus found represent the common characterictics of the data. In order to quantify the difference among the subjects with respect to the factors at hand, factor scores for each subject are computed.

Factor scores define the position of each subject in the common factor space. In our example, each country will have a score per factor. Factor scores can be defined as follows:

(12) $$\hat{F}_{ij} = B_1 X_{i1} + B_2 X_{i2} + ... + B_p X_{ip}$$

where $i = 1,...,n$ refers to subjects (the countries in our example), $j = 1,...,q$, refers to the factors, and X_{il} is the value of variable l for subject i.

Equation (12) can be written more compactly as:

(13) $$\underset{(nxq)}{\hat{F}} = XB$$

If we define z as the standardized from of matrix X , it can be shown that the factor score matrix is given by:[15]

(14) $$\hat{F} = zR^{-1}\Lambda$$

In this paper, I will use factor analysis to underline the common characteristics underlying the data at hand, and use factor scores in order to present yet another overall classification of Arab countries.

3. The Data

In order to operate overall as well as sector-specific classifications of Arab countries, data on several socio-economic indicators were collected for three reference years: 1970, 1980, and 1992. The classification of Arab countries over the three years is intended to describe how Arab economies' characteristics have evolved over the last three decades or so.

The variables used were classified according to broad criteria into : Labor Market, Economic, Fiscal, Financial, External, Population and Demographic, and Social variables.

Table 1 presents the catalogue of variables used in the classification, Table 2 provides the definitions of these variables, and Table 3 provides the main sources of the data used.[16] The number of variables as well as the countries used were based on data availability. The list of countries included in the analysis is provided in table 4.

[15] See, for instance, Dillon and Goldstein (1984) for the proof.

[16] The main data for this study and their original sources are derived form Dahel (1995), Wadie (1995). A big part of the data was also provided thanks to the efficient research assistance of Salih Al-Asfour, Jamal Hamed, and Hadeel Abu Loghod, from API.

TABLE 1. : CATALOGUE OF VARIABLES

	1970	1980	1992
Labor Market	TLF	TLF	TLF
	NEMR	UR	LF/POP
	ALFR	NEMR	NEMR
	WLFPRT	ALFR	ALFR
	WLFPR	ILC	ILC
	LFPRT	WLFPRT	WLFPRT
	LFPR	WLFPR	WLFPR
	ILFR	FPRT	LFPRT
		LFPR	LFPR
		ILFR	ILFR
Economic	AGDP	AGDP	AGDP
	IGDP	IGDP	IGDP
	CGDP	CGDP	CGDP
	SGDP	SGDP	SGDP
	PCX GDP	XCR	XCR
	GDP/C	GDP/C	PCXGDP
	I	I	GDP/C
	PC	PC	I
		INFLATION	PC
			INFLATION
Fiscal Sector	CurExp	G/R	G/R
	G/R	FDY	FDY
	FDY	GC	GC
	GC	TAXREV	TAXREV
	TAX REV	EER	EER
		SER	SER
		SSER	SSER
Financial Sector	DCP	DCP	DCP
	NFA	NFA	NFA
	S/GDP	S/GDP	FSM
	M2/GDP	M2/GDP	S/GDP
			M2/GDP
External Sector		**Current Account Indicators**	**Current Account Indicators**
	DOO	DOO	DOO
	CAB/GDP	CAB/GDP	CAB/GDP
	FUELX	TB/GDP	TB/GDP
	MANUFX	FUELX	FUELX
	FUELI	MANUFX	MANUFX
	MANUFI	FUELI	FUELI
	DEBT/GDP	NWR	MANUFI
	DEBT/X	MANUFI	
		Capital and B.O.P Indicators	**Capital and B.O.P Indicators**
		BOP/GDP	BOP/GDP
		IR	IR
		ERR	DEBT/GDP
			ERR
			FDI/INF
			DEBT/X
Population and Demographic	POP	POP	POP
	UPOP	UPOP	UPOP
	TFR	TFR	TFR
	CBR	CBR	CBR
	CDR	CDR	CDR
	IMR	IMR	IMR
Social	LE	LE	LE
	POPH	POPHD	POPHD
	POPHB	POPHB	DR
	E/GNP	E/GNP	POPHB
	IR	TR	E/GNP
	DR	DR	IR
			ERE
Total Number of Variables	**45**	**53**	**58**

TABLE 2. : VARIABLES DEFINITION

Labor Market Variables

TLF:	Total Labor Force.
UR:	Unemployment Rate.
LF/POP:	Labor Force as a percentage of Total Population.
NEMR:	Net Emigration Rate.
ALFR:	Percentage of Labor force in Agriculture.
ILC:	Labor Cost in Industry.
WLFPRT:	Women Labor Force Participation Rate (Ratio of Female in labor force to female population).
WLFPR:	Women Labor Force Participation Ratio (share in Tolal labor force).
LFPRT:	Labor Force Participation Rate.
LFPR:	Labour force Participation Rate.
ILFR:	Percentage of Labor Force in Industry.

Social Variables

LE:	Life Expectancy at Birth.
POPH:	Population per Physician.
POPHB:	Population per Hospital Bed.
E/GNP:	Education Expenditure as pertencage of GNP.
DR:	Dependency Ratio.
IR:	Illeteracy rate among Population above 15 years.

Population and Demographic Variables

POP:	Population
UPOP:	Urban Population as Percentage of Total
TFR:	Total Fetility Rate
CBR:	Crude Birth Rate
CDR:	Crude Death Rate
IMR:	Infant Mortality Rate

External Sector Variables:

DOO:	Degree of Openness (Exports plus imports of goods and services over GDP).
CAB / GDP:	Current Account Balance as a Percentage of GDP.
TB / GDP:	Trade Balance as a Percentage of GDP.
FUELX:	Fuel Exports as a Percentage of Total Merchandise Exports.
MANUFX:	Manufactured Exports as a Percentage of Total Merchandise Exports.
FUELI:	Fuel Imports as a Percentage of Total Merchandise Imports.
MANUFI:	Manufactured Imports as a Percentage of Total Merchandise Imports.
NWR:	Net worker's Remittances as a percentage of Exports.
DEBT / GDP:	Debt as a percentage of GDP.
DEBT / X:	Debt Service Ratio.
BOP / GDP:	Balance of Payments as a Percentage of GDP.
IR:	International Reserves in millions of US Dollars.
ERR:	Exchange Rate Regime[*] .
FDIINF:	Foreign Direct Investment Inflows.

Economic Variables:

AGDP:	Share of Agriculture in GDP.
IGDP:	Share of Industry in GDP.
CGDP:	Share of Construction Sector in GDP.
SGDP:	Share of Service Sector in GDP.
XCR:	Export Concentration Ratio.
PCXGDP:	Share of Primary Commodity Exports in GDP including Fuel.
GDP / C:	GDP per Capita.
I:	Investment Ratio (Share in GDP).
PC:	Personal Consumption Ratio (Share in GDP).
INFLATION:	Rate of Growth of Consumer Price Index.

Fiscal Variables:

CUREXP:	Share of Current Expenditures in total Government Expenditures.
G/R:	The Ratio of Total Expenditures to Total Revenues.
FDY:	Fiscal Deficit as a percentage of GDP.
GC:	Ratio of Government Consumption to GDP.
TAXREV:	Share of Tax Revenues in total Government Revenues.
EER:	Share of economic expenditures in total Government expenditures.
SER:	Share of social services expenditures in total Government expenditures.
SSER:	Share of general service expenditures in total Government expenditures.

Financial Variables:

DCP:	Share of Net Domestic Credit to the Private Sector in Total Credit.
NFA:	Net Foreign Assets of the Banking System (Millions of Dollars).
FSM:	Number of firms quoted in the stock market.
S/GDP:	Domestic Saving Ratio.
M2 / GDP:	M2 as a share in GDP (Financial Deepening Indicator).

[*] This is a discrete variable defined as follows: ERR=5 (independent float); ERR=4 (managed floating); ERR=3 (Peg to a Basket); ERR=2 (peg to SDRs or US$); ERR=1 (fixed with respect to US $)

TABLE 3. : MAIN SOURCES OF DATA

Type / Variable Name	Sources
Labor Market	
TLF	Social Indicators of Development (SID), and World Tables (WT), the World Bank
LF/POP	Claiming the Future, World Bank (1995)
UR	SID, International Labor Organization (ILO)
NEMR	Computed
ALFR	Human Development Report (HDR), UNDP, Unified Arab Economic Report (UAER)
ILC	UAER, Arab Organization for Industrial and Mining Development, ESCWA.
LFPR	SID
LFRPT	SID
WLFPR	SID, WT
WLFRRT	SID, HDR
ILFR	HDR, UAER
Economic	
AGDP	AMF, General Secretariat of the Arab League, UAER, ESCWA
IGDP	AMF, General Secretariat of the Arab League, UAER, ESCWA
CGDP	AMF, General Secretariat of the Arab League, UAER, ESCWA
SGDP	AMF, General Secretariat of the Arab League, UAER, ESCWA
XCR	Economic Trends in the MENA Region, Economic Research Forum (ERF), 1996.
PCXGDP	ERF, 1996.
GDP/C	Arab Monetary Fund (AMF), UAER, Arab Economic Unity Council (AEUC),
I	AMF, ESCWA, General secretariat of the Arab League, UAER
PC	AMF, General Secretariat of the Arab League, UAER, and other Sources
INFLATION	WT, OAPEC, and other various sources.
Fiscal Sector	
CurExp	Computed (AMF, ESCWA, and other sources)
G/R	Computed (AMF, ESCWA, and other sources)
FDY	Computed
GC	AMF, General Secretariat of the Arab League, UAER, and other Sources
TAXREV	AEUC, ESCWA, and other Sources
EER	Arab Economic Unity Council, UAER, General Secretariat of the Arab League
SER	Arab Economic Unity Council, UAER, General Secretariat of the Arab League
SSER	Arab Economic Unity Council, UAER, General Secretariat of the Arab League
Financial Sector	
DCP	AMF
NFA	AMF
S/GDP	Computed (AMF)
M2/GDP	Computed
FSM	AMF and other sources
External Sector	
DOO	Computed
CAB/GDP	Balance of Payments Statistics (BOPS) International Financial Statistics (IFS), AMF
TB/GDP	Balance of Payments Statistics (BOPS) International Financial Statistics (IFS), AMF
FUELX	WT, BOPS
MANUFX	WT, BOPS
FUELI	UNCTAD
MANUFI	UNCTAD
DEBT/GDP	WT
DEBT/X	WT
NWR	WT, BOPS, UAER
BOP/GDP	BOPS, IFS, AMF
IR	WT, BOPS, UAER
ERR	IMF
FDIINF	ERF, 1996
Population and Demographic	
POP	SID
UPOP	SID
TFR	SID
CBR	SID,UN Statsistical Yearbook
CDR	SID, UN Demographic Yearbook
IMR	SID, UN Demographic Yearbook, UAER
Social	
LE	SID, UAER
POPH	WDR, World Bank, UAER
POPHB	General Secretariat of the Arab League
E/GNP	UNESCO, SID, WT, HDR
DR	Computed (Based on WDR and UAER)
IR	SID, HDR, UNESCO, UAER

TABLE 4. : CATALOGUE OF COUNTRIES

1970	1980	1992
Algeria	Algeria	Algeria
Bahrain	Bahrain	Bahrain
Egypt	Egypt	Egypt
Iraq	Jordan	Jordan
Jordan	Kuwait	Kuwait
Kuwait	Libya	Lebanon
Lebanon	Mauritania	Libya
Libya	Morocco	Mauritania
Mauritania	Oman	Morocco
Morocco	Qatar	Oman
Oman	Saudi Arabia	Qatar
Qatar	Sudan	Saudi Arabia
Saudi Arabia	Syria	Sudan
Sudan	Tunisia	Syria
Syria	UAE	Tunisia
Tunisia		UAE
UAE		Yemen
Yemen		
TOTAL		
18	15	17

4. THE RESULTS

4.1. General Classification of Arab Countries

In this section, Cluster and Factor Analyses, respectively, are used to adopt an overall classification of Arab countries according to their socio-economic characteristics. The classification was performed for the three periods 1970, 1980, and 1992, in order to assess whether the characteristics of these countries have changed.

Cluster Analysis Classification :

Using standardized data, the four classifications methods underlined earlier namely, Complete Linkage, Single Linkage, Ward, and K-Means methods, were used to classify Arab countries according to their economic characteristics. The first three methods were implemented using two different distance measures: Euclidean and Pearson. The K-Means method uses, by its very definition, an Euclidean type of distance.

A relative stability of the classifications, reported in appendix one, was striking. The general classification of Arab countries into oil and non-oil producing countries was found to be the most stable and regular classification. This classification was stable not only over time but also under different classification methods and distance measures. This clearly shows that oil and non-oil producing countries tend in general to share several of the socio-economic characteristics underlined in this paper and therefore is an appropriate classification for general types of analysis.

A finer classification of these two groups, over the three periods considered, has revealed other stable taxonomies. For instance, for the 1992 classification, and for oil producing countries, it was found that the six countries of the Gulf Cooperation Council (GCC) namely, Saudi Arabia, Kuwait, United Arab Emirates (UAE), Bahrain, Qatar, and Oman, tend to form a separate cluster especially if correlation among variables is taken into account. However, classification results reveal that Saudi Arabia and Oman are closer to the subgroup of oil countries formed by Libya and Algeria than the remaining four GCC countries. In fact, the last four countries tend to systematically cluster in the same group. A closer look at the group of the four GCC countries, shows that Qatar and UAE form more homogenous a group than any other pair. Kuwait and Bahrain, although belonging to the same group most of the time, are forming, each, its own group if finer classification is adopted.

With respect to the non-oil country group, a natural classification by income level has been found. Mauritania, Sudan, Yemen, and Egypt constitute the group of low-income Arab countries in the sample. Morocco, Tunisia, Jordan, Syria and Lebanon is the other group of medium-level income. The results also show a relatively stable pairing of countries. For instance, Morocco and Tunisia on one hand, and Jordan and Syria on the other hand, tend to form homogeneous groups if a finer classification is adopted.

Similarly, Qatar and UAE, and to a lesser extent, Algeria and Libya, tend to form regular pairs in the group of oil producing countries. The general classification of countries has also revealed that certain countries are dissimilar to the countries of the group they belong to. The most striking cases are that of Egypt, and to a lesser extent, Lebanon and Kuwait.[17]

The results show that, over time, few countries have moved closer to one another, while for others the exact opposite happened. The 1980 classification was pretty much similar to the 1992 classification. The only difference being that Algeria was not part of the oil-country group. Saudi Arabia was found to cluster with Libya, and to a lesser extent with Oman. Egypt exhibited characteristics that distanced it from the low-income group of non-oil producing countries.

The 1970 classification was pretty much the same as the previous ones except that few countries looked closer together than they actually appear to be. Bahrain and Kuwait, for instance, looked more like a pair of homogeneous countries.

Factor Analysis Classification :

Using the data at hand, Factor Analysis was applied, first to underline, for each period considered, the main unobservable factors accounting for most of the variation in the variables at hand, related to the Arab countries, and second to use the scores of each country on the latter factors in order to perform an overall classification of these countries.

Classification of countries based on factor scores was adopted using the Complete Linkage Method and the Euclidean and Pearson distance measures. The results of this classification are reported in appendix two.

[17] For the case of Kuwait, the state of the economy has been affected dramatically by the effects of Iraqi invasion of Kuwait and the aftermath of the second Gulf War. Similarly, the civil war is in part responsible for the peculiar characteristics of Lebanon.

Generally speaking, five important factors seem to account for more than seventy percent of the variation among Arab countries with respect to the variables included. The first factor could be named the “Development Factor” and account for the three periods considered for over 30 percent of total variation. The Development factor includes variables that are generally associated with the development level of each country. These include for instance, the GDP per capita, share of agriculture in GDP, share of tax revenues in total Government revenues, degree of openness, share of manufacturing in total merchandise imports, level of urbanization, crude birth and death rates, level of indebtedness, and the share of women in total labor force.

The second factor can be labeled the “External Factor”. It accounts for over 15 percent of total variation. Many variables such as those related to the current and balance of payments accounts, manufacturing exports, and level of international reserves, load into this factor.

The three remaining factors vary in importance according to the period considered. These factors were, “Population-Labor Factor”, “Fiscal-Monetary Factor” and “General Economic Factor”. The last factor is a blend of a variety of economic indicators.

The number of factors retained in the analysis was determined by their relative importance in accounting for the total variation of the variables at hand, and the number of variables loading into these factors.

For 1992 and 1970 the five factors were retained. For 1980, the fifth factor namely, the General Economic Factor, was dropped because only one variable was loading into this factor.

Classification of countries based on factor scores was adopted using the three hierarchical classification methods adopted above, and the Euclidean and Pearson distance measures. The results of this classification are reported in appendix two.

Broadly speaking, the classification based on Factor Analysis confirms to a great extent the results of the classification based on Cluster Analysis. Oil-non-oil classification seems again to be the most relevant criterion to classify Arab countries. On the other hand, oil countries were found to form a more homogeneous and stable group than non-oil countries.

However, the classification based on Factor Analysis points again to the fact that oil-countries are less homogeneous a group than what is believed to be. The classification shows that Saudi Arabia, Libya, and Oman are closer to each other than to the group made of Bahrain, Kuwait, UAE, and Qatar. Moreover, the last two countries, are consistently belonging to the same group. The remaining Arab oil-producing countries, Algeria and Iraq, represent border cases. Their distribution over clusters is somewhat unstable over time and across clustering methods.

In the non-oil countries group, there are less stable patterns in classification. There is however a general homogeneity between low-income countries such as Sudan, Mauritania, and Yemen. On the other hand, two pairs of countries seem to be more homogeneous than others. The first pair comprises Tunisia and Morocco and the second, Jordan and Syria.

Finally, few countries have peculiar characteristics and could hardly pair with any other country. As pointed out earlier, Egypt, Kuwait, and Lebanon, represent typical examples in this regard.

4.2. Sector-Specific Classification of Arab Countries

The overall classification of countries is useful in facilitating general types of analysis. This type of classification is, however, not appropriate to conduct finer analysis at the sector or micro level. The rest of this section analyzes the sector-specific classification results, related to the three periods adopted and reported in appendix three. Classification was performed using the Complete Linkage Method and the Euclidean as well as the Pearson distance measures.

Labor Market Classification

The 1992 Labor market classification reveals that Arab countries tend to be naturally split between labor-importing and labor-exporting countries especially if the Pearson distance measure is used. The labor exporting countries in turn can be divided into large-labor exporting and small-labor exporting countries. Large and small refer here to the size of labor emigration. The large labor-exporting group comprises Morocco, Sudan, Yemen, Syria, Algeria and Egypt, while the low labor-exporting group comprises Lebanon, Tunisia, and Mauritania.

Labor importing countries are made of mainly oil-producing countries including the six GCC countries and Libya. The only exception in this group is that of Jordan. Jordan membership to this group can be explained by the coincidental similarity of its labor market characteristics to the rest of the countries in this group notably in terms of the size of the labor force and its relative importance, the share of labor force in agriculture and industry, the degree of participation of women in the labor force, and the labor cost in the industrial sector. In the labor-importing groups, four of the GCC countries namely, Bahrain, Kuwait, Qatar, and the UAE share closer characteristics than the rest of the countries in the group.

The classification results also show pairing of several countries such as Morocco and Sudan, Qatar and UAE , and Jordan and Libya.

The 1980 classification shows a relative stability of some of the main results of the previous classification. The most noteworthy, are the classification of Arab labor markets into labor-exporting and labor-importing, and the stability of the group of the four GCC countries: Bahrain, Kuwait, Qatar, and UAE . The pairing of countries was, however, somewhat different from what the previous classification has suggested.

The 1970 classification is not very different from previous classifications. Most of the oil-producing countries still represent a distinguishable group from the rest of the countries. However, the contrast between labor-exporting and labor importing countries is not as pronounced as it turns out to be for the 1980 and 1992 classifications. This can be explained by the fact that the positive effects of oil-shocks have been a major driving force behind the movement of Arab labor. The salient features of the 1970 classification are the relative similarity of labor characteristics in Saudi Arabia, Yemen, and Oman, on the one hand, and Qatar and UAE, on the other.

Economic Classification

The economic classification of Arab countries over the three periods under study shows clearly the importance of the oil sector in such classification. Oil producing countries tend to have similar economic structures notably, in terms of the share of the different sectors in GDP, the degree of export concentration, the share of primary commodity exports in GDP, inflation and saving rates, and investment ratios.

In the 1992 classification, the Arab countries are naturally divided into two major groups : oil-producing countries and non-oil producing countries. Furthermore, the non-oil producing countries are divided into a low-income group comprising Syria, Yemen, Mauritania, and Sudan; and medium-income group comprising Egypt, Morocco, Tunisia, Jordan, and Lebanon.

In the oil-producing group, the six GCC countries appear more homogeneous a group than the remaining two Arab oil producing countries included in the 1992 sample, namely, Algeria and Libya. The latter countries tend to cluster together.

The 1980 classification, reveals also a division of countries between oil and non-oil countries. The oil group comprises all six GCC countries plus Libya, while the non-oil group is made of the rest of the countries. The same remarks apply for 1970. In addition, the non-oil group is divided into a low-income group comprising Sudan, Yemen, Mauritania and Iraq, and a group of medium-income and diversified economies including Morocco, Syria, Tunisia, Jordan, Lebanon. Egypt is a border case between the last two groups and Algeria is not very similar to the rest of the countries in the medium-income group.

Fiscal Sector Classification

Unlike previous classifications, there seem to be no general features that confer a relative stability of the fiscal sector classification across periods and clustering methods. This is said, oil countries tend share some of the characteristics notably, the low share of tax revenues, large fiscal deficits, and high share of economic expenditures.

The 1992 classification shows that many oil-producing countries seem to have more common characteristics than the rest of the countries. However, several oil-producing countries such as Bahrain and UAE can hardly be classified within the same group as that of the rest of the GCC countries. On the other hand, a country like, Kuwait, which has obviously been affected by the Iraqi invasion and its aftermath, looks different from the rest of the GCC countries.

Small diversified economies like Lebanon, Tunisia and Jordan seem to share common characteristics. The rest of the results show the clustering of countries that look very different in other characteristics. This points to the fact that their similarity in this front is only coincidental.

The 1980 classification shows a more pronounced heterogeneity of Arab countries. Only few oil countries such as UAE, Kuwait, and Bahrain maintained a relatively similar fiscal structure. In the non-oil producing countries, only Egypt, Jordan, and Sudan appear to have close fiscal indicators.

For 1970, the fiscal structure seems to be affected by whether the country is a major oil-producing or not. Saudi Arabia seems to be somewhat different from the rest of the group. On the other hand, large countries including Iraq, Algeria, Sudan and Egypt seem to have closer fiscal characteristics. These countries have in common a high level of Government expenditures and over-taxed economies. The rest of the countries are hardly making a homogeneous group from a fiscal perspective.

Financial Sector Classification

The classification of Arab countries over the three periods considered shows that the financial structure is somewhat influenced by whether a country is a major oil-producer. Oil revenues have enabled oil-producing countries to expand their GNP per capita and accumulate large amounts of foreign assets. This has resulted in high saving rates that necessitated the development of a modern and efficient financial sector.

The 1992 classification has not revealed any obvious or stable classification of Arab countries. This points to the fact that Arab countries are too heterogeneous to form meaningful clusters. The "oil-effect" mentioned above shows clearly in the 1980 classification where the financial structure seems to be polarized between oil and non-oil countries. GCC countries tend to form a uniform group except for Saudi Arabia which exhibits special financial indicators. The non-oil countries could not be classified into stable and homogeneous sub-groups.

The 1970 classification shows that the financial indicators of Arab countries are quite different. However, oil producing countries have remarkably higher scores on their financial indicators especially in terms of saving rates and foreign financial assets. Lebanon and Jordan have, on the other hand, a high level of financial deepening as depicted by the M2/GDP ratios.

External Sector Classification

Overall, oil-non-oil criterion matters to a great extent in the classification of Arab countries according to their external sector. The 1992 and 1980 classifications show that Arab countries are clearly divided into oil and non-oil countries although this classification is more obvious for the case of the current account indicators than that of capital and balance of payments indicators.

Oil countries tend to exhibit a high degree of openness, high shares of fuel exports and manufacturing goods imports, large surpluses (small deficits) in their current account and balance of payments deficits, large stocks of international reserves, low levels of international indebtedness, and a fixed-type of exchange rate regimes. Even for the 1970 classification where, for lack of data, the classification of

indicators into current and capital and balance of payments indicators is not adopted, oil producing countries tend to exhibit discernible common external sector characteristics.

In contrast, non-oil countries have very few discernible common characteristics which consists of having persistent current and balance of payments accounts, and being debt ridden. The more diversified economies such as Morocco, Tunisia, Syria, Egypt and Jordan, tend to have high shares of manufacturing in exports as well as imports.

Population and Demographic Classification

The classification of Arab countries based on their population and demographic characteristics, show a relative instability of this classification over the last three decades or so.

In the 1990's, there is a clear classification of Arab countries into relatively homogeneous groups. Four groups are distinguishable. Group one is made of three populous Arab countries namely Algeria, Morocco and Egypt with a rate of urbanization close to 50 per cent and close birth, death and infant mortality figures. The second group comprises small size countries with a fairly high level of urbanization and close demographic indicators. The countries of this group are : Bahrain, Qatar, UAE, Kuwait, Tunisia, and Lebanon. The third group is made of low-income Arab countries namely, Mauritania, Yemen, and Sudan with low level of urbanization and high birth, death, and infant mortality rates. The fourth group comprises, Saudi Arabia, Syria, Jordan, Libya and Oman. These countries have fairly high fertility and birth rates and low death rates.

In the 1980's, Arab countries can also be classified into four main groups. The first group comprises four out of the six GCC countries namely, Bahrain, Kuwait, Qatar and UAE. These countries are not very populous and are characterized by a high degree of urbanization and have similar demographic characteristics. The second group is made of Morocco, Tunisia, and Egypt which have close urbanization rates (less than, or equal to, fifty percent) and share the same demographic characteristics. The third group is made of Saudi Arabia, Syria, Algeria, Jordan, and Libya. These countries have relatively high urbanization rates and crude-birth to-death rate ratios. The last group comprises Mauritania, Sudan, and Oman. These countries are characterized by their low urbanization rates and high fertility, birth, and death rates.

In the 1970's, the population characteristics of Arab countries look somewhat different than previously given the redistribution of countries across clusters. Few features are noticeable. Low-income countries such as Sudan, Mauritania and Yemen seem to share many demographic and population characteristics. The other noticeable factor is the pairing of several countries such as, Qatar and UAE, Algeria and Morocco, Syria and Libya, and Oman and Yemen.

Social Sector Classification

Classification of Arab countries according to their social indicators shows the importance of income as a determinant of social performance especially for high and low income groups.

In the early 1990's, the classification of Arab countries according to their social indicators shows that, except for the countries with the lowest scores, it is not very sensitive to the level of income. Countries like Tunisia, Syria, Lebanon, and Jordan, which have managed to maintain a good balance between their income level and social development, are found to exhibit pretty closer social indicators to high-income countries such as Kuwait, UAE, and Qatar.

The classifications of the 1980's and 1970's, reveal the importance of the level of income in these classifications. In the 1980's classification, the high-score group is made of oil-producing countries with high level of per capita income such as Kuwait, UAE, Bahrain, and Qatar. The low-score group is mainly made of countries with low levels of income.

In the 1970's, the classification of Arab countries according to their social indicators reveal the presence of three major groups. The first group of low social performers is made of, as indicated, of low-income countries such as Sudan, Yemen, Mauritania, and Oman and to a lesser extent Morocco. The second group involves countries with higher level of income than the first group. These include Algeria, Saudi Arabia, Iraq, Tunisia, Egypt, Libya, Syria, and Jordan. Many countries in this group have higher scores than is warranted by their level of income. The third group includes countries with the highest level of income in the region namely, UAE, Kuwait, and Qatar. The group comprises also Lebanon which managed to maintain high scores of social indicators in spite of its below group-average level of income.

5. CONCLUDING REMARKS

It has been shown in this paper that the general classification of Arab countries into oil and non-oil producing countries is not only justifiable but represents the most, if not the only, meaningful criterion to adopt an overall classification of these countries according to their socio-economic characteristics.

The overall classifications of Arab countries based on a wide variety of socio-economic variables, performed with the statistical classification methods of Cluster and Exploratory Factor Analyses respectively, and over the three periods of 1970, 1980, and 1992, are broadly in line with each other. They point to the fact that the criterion of "oil-versus-non-oil" gives the most stable and regular classification of Arab countries. These classifications were stable not only over time but also under different classification methods and distance measures.

The results of Factor Analysis have revealed that five important factors account for most of the variation among Arab countries with respect to the variables used in the classification. These are the "Development Factor" including variables related to the development level of a given country; the "External Factor" including variables related to the current account and the balance of payments, the "Population-Labor factor"; the "Fiscal-Monetary Factor"; and the "General Economic Factor".

Aside from the oil-non-oil criterion, the general classification of Arab countries, has revealed many other interesting features. First, the six GCC countries represent the most homogeneous group among Arab countries. Second, within the latter group, Bahrain, Kuwait, Qatar, and the UAE are more similar than the two remaining GCC countries namely, Oman and Saudi Arabia. Third, within the sub-group of the four GCC countries, Qatar and the UAE share more common characteristics than the two remaining countries. Fourth, in the non-oil producing group, low-income countries such as Mauritania, Sudan, and Yemen form a more homogeneous group than the rest of the non-oil countries. Fifth, countries with diversified economies such as, Egypt, Jordan, Lebanon, Morocco, Syria, and Tunisia, tend to have several characteristics in common.

The sector-specific classification of Arab countries has provided yet further evidence of the importance of the "oil criterion" in the classification of these countries. Oil-producing countries have been found to share similar economic structures, external, fiscal, and financial sectors indicators.

The economic structure of these countries is dominated by the oil sector and oil-related industries. The low level of diversification and the dominance of the oil sector in production, have resulted in high export concentration ratios in favor of fuel-related products. Given the favorable world oil market conditions that prevailed for a relatively long period, most Arab oil-producing countries have had current account and balance of payments surpluses that allowed them to accumulate large portfolios of foreign assets.

The substantial oil revenues have enabled these countries to pay their large import bills of food, equipment, and machinery. These revenues have also contributed to the financing of large infrastructure and oil-related investment projects and allowed oil-producing countries to follow lax fiscal policies resulting in large budget deficits.

The oil market boom during the 1970's and early 1980's, has resulted in a substantial increase in income per capita in the oil-producing countries. This in turn has led to the increase in saving capacity that necessitated the development of a modern financial market.

Even in the sectors that are not directly affected by oil , the results show that the oil criterion still has a bearing on classification outcomes. Labor market classification shows that the most relevant criterion for classification is that between labor-exporting and labor-importing countries. However, labor importing countries are mostly oil-producing countries, whose oil-wealth has enabled them to absorb labor surpluses from non-oil producing countries over the last twenty years or so.

On the other hand, large oil-producing countries notably, the GCC countries, have improved the state of their social development by investing in health, education, and human development. This explains their higher ranking in terms of the HDI index and their similar scores with respect to the social variables included in this paper.

The classification of Arab countries according to population and demographic characteristics, is perhaps the only sector-specific classification where the "oil-effect" has not been found to matter. In fact, the population and demographic classification results turned out to be the most unstable over time and clustering methods.

For the rest of the countries, the different sector-specific classifications have not indicated any stable groupings of these countries neither over the years nor clustering methods. The most possible relevant criteria are the level of income and the degree of economic diversification.

Income has been found to matter, to some extent, only in the economic, population and demographic, and social classifications of the latter countries. Diversified economies tend also to share common economic and external sector structures.

The outcomes of this paper can be extended and improved in several directions. However, data, which have been the most constraining factor in the study, is the most noteworthy direction for improvement.

The use of multivariate statistical analysis such as, Cluster and Factor Analyses, require that data for all the countries and all the variables be complete. Missing data procedures in this case are neither well defined nor agreed upon. If the data problem can be overcome, the introduction of other variables related to the sectors included in the paper (notably, those related to the labor market and financial sector) or other additional sectors can be very insightful in providing other criteria for classifying Arab countries. Few additional sectors are of immediate relevance given the recent developments in the international arena. These include, the environment, institutions, governance factors, and poverty.

REFERENCES

References

Al-Kawaz, A. (1995); **"Economic Classification of Arab Countries for Building a Prototype Policy Evaluation Model"**, (Arabic), Research Proposal, Arab Planning Institute - Kuwait.

Amrouch A. and A. Wadie (1987); **"Analysis of Socio-Economic Characteristics of Arab Countries"**, (Arabic), Arab League Educational, Cultural, Science Organization.

Arab League General Secretariat et al.; **"The Arab Unified Economic Report"**, various issues.

Dahel, R. (1995); **"Foreign Trade, Exchange Rate and Balance of Payments-Characteristics of Arab Economies"**, Arab Planning Institute - Kuwait.

Dellaportas, G. (1983**); "Classification of Nations as Developed and Less Developed : An Arrangement by Discriminant Analysis of Socio Economic Data"**, *American Journal of Economics and Sociology,* Vol. 42, No. 2, April, pp. 153-66.

Dillon, W.R. and M. Goldstein (1984); **"Multivariate Analysis : Methods and Applications"**, John Wiley & Sons, New York.

Everitt, B.S. (1993); **"Cluster Analysis"**, third edition, Edward Arnolds, London.

Gorsuch, R.L. (1983); **"Factor Analysis"**, second edition, Lawrence Elbaum Associates Publishers, Hillsdale, New Jersey.

Hassib, K. et al. (1988); **"Foreseeing the Future of Arab Countries : Final Report"**, (Arabic), Centre for Arab Unity Studies, Beirut, Lebanon.

Organization of Arab Petroleum Exporting Countries (OAPEC) and Enter Nazionale Idrocarburi (ENI) (1985); **"The Interdependence Project"**, various documents.

Saadeddine, E. et al. (1989); **"Foreseeing the Future of Arab Countries Project : Arab Countries Development"**, (Arabic), Centre for Arab Unity Studies, Beirut, Lebanon.

United Nations Development Program; **"Human Development Report"**, various issues, Oxford University Press.

Vogel, F. (1993); **"Some Remarks on a Classification of the Countries of the World According to their Stage of Development"**, *Jahrbucher-fur-Nationaloknomie-und-statistic*, G.F. Verlag Stuttgart, Vol. 211, part 3-4, pp. 306-23.

Wadie, A. (1995); **"Human Resources : Population and Labor Market Characteristics in the Arab Countries"**, (Arabic), Arab Planning Institute - Kuwait.

World Bank (1996); **"World Development Report"**, Oxford University Press.